AMENDMENTS

Amendments to the Specification:

Please replace paragraph [004] with the following rewritten paragraph.

[004] Volumes of material come into and are stored in archives by television, radio, and news production facilities today. In a previous technology, ~~personal~~ personnel manually generated software tags for the content in this material. The software tags contained descriptors about the audio/video data in order to assist searching through and finding a desired piece of information in the volumes of material. Journalists, interns or researchers listen to hours of tape manually searching and analyzing through the recorded information to find the exact segment or piece of knowledge that the person was seeking. Limited sets of audio content were tagged because the manual process of tagging is expensive. Additionally, the non-standardized methods for tag coding generate high error rates during the search process.

Please replace paragraph [0018] with the following rewritten paragraph.

[0018] Two exemplary XML documents **202, 204** produced by attribute filters will be described. An XML document has a hierarchical layered structure including tag-value pairs and elements of that tag value pair. In an embodiment, an XML tag value pair **205** defines the particular type of attribute, such as an Accent change **208**, an identified word **206**, a human language being spoken, etc. and elements **207** of the particular tag value pair **205** define characteristics of that attribute, such as a time code **210**, etc. For example, a first tag-value pair **206** may contain the identified attribute of "Utterance-

Word" which indicates the attribute filter detected an Utterance and identified the utterance as a word. Similarly, a second tag-value pair **208** may contain the identified attribute of "Utterance-Accent Change" which indicates that the attribute filter detected ~~an~~ a change of accent in the spoken words and identified the time frame during which that accent was being used.

Please replace paragraphs [0022] and [0023] with the following rewritten paragraphs.

[0022] In an embodiment, the index control module **104** generates the time codes which each attribute filter **108, 109, 110, 112, 114** attaches as an element for each identified attribute. Therefore, every identified attribute of an information stream may be referenced to the same time keeping mechanism, such as a clock or video frame counter. In an embodiment, the index control module **104** generates time codes for each identified attribute referenced to the time indication or the video frame count. In an embodiment, the index control module **104** uses a time ordered index via a data table that keeps track of the data by a time indication and may be correlated by the time indication. In an embodiment, an integrated time ordered index consists of some or all of the attributes identified by the ~~attributes~~ attribute filters from the same information stream merged into a single index of identified attributes all sharing a common time reference.

[0023] In an embodiment, a first attribute from the information stream may be correlated or ~~attributes~~ paired with other attributes having a similar time code. The similar time code may be a first attribute possessing the same start time code as a second attribute or due to the duration of the attributes overlapping with each other. For

example, separate identified attributes of the same information ~~stream~~ stream, such as a video clip of a person speaking and the ~~transcripted~~ transcribed text of the person's speech, may be frame accurate because they are both referenced to the same synchronized time code. Thus, the sequence of time coded XML documents containing a video track from pan to fade or scene change to scene change may be synchronized with an audio track corresponding to the audio XML documents having the same sequence of time codes.

Please replace paragraphs [0054] through [0056] with the following rewritten paragraphs.

[0054] For each word, the attribute filter generates an individual XML document including as elements the identified word, the confidence rating from each language model, and the time code for that word. The transcript of the entire conversation corresponds with an overall XML document for that conversation. However, because each word is a discreet XML document itself within the aggregate XML document, then a user may select a phrase or portion within the transcript and start playing the audio segment from that exact moment in time corresponding to the selected phrase.

[0055] In an embodiment, the audio information stream **310** passes through the attribute filter in a single pass rather then the audio information stream being sequentially analyzed first by a first language model and then by the next language model. In an embodiment, generating an integrated index including the data from the four language models in a single pass assists processing an audio information stream in real time. In an embodiment, an integrated time ordered index consists of some or all of the attributes

identified by the ~~attributes~~ attribute filters from the same information stream merged into single index of identified attributes all sharing a common time reference.

[0056] Referring to figure 1, in which the manipulation-module **124** interacts with the storage devices **116** and the intelligence engine **120** to allow a user to navigate and utilize the indexed information stream data. Due to the identified attributes ~~begin~~ being organized through a time ordered index, transmodal manipulations of each type of attribute may occur. A user through the user interface may perform operations on a first set of attributes in order to manipulate a second set of attributes. For example, a user may create a new audio clip of a desired segment of a radio broadcast by highlighting the ~~transcripted~~ transcribed text and cutting the transcript text from the text document. Further, the user may splice multiple video clips together by assembling and inserting text corresponding to each video clip. Thus, the user manipulates a first type of attribute such as the transcripted text in order to perform an operation on the second type of attribute such as spoken words or video characteristics.